

Tenant-based measured electricity use in 4 large office buildings in Tallinn, Estonia

Andrea Ferrantelli^{1*}, Hans Kristjan Aljas¹, Vahur Maask², Martin Thalfeldt^{1,3}

¹Department of Civil Engineering and Architecture, Tallinn University of Technology, 19086 Tallinn, Estonia

²Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, 19086 Tallinn, Estonia

³Smart City Center of Excellence, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia

Abstract. The energy performance assessment of buildings during design is usually based on energy simulations with pre-defined input data from standards and legislations. Typically, the internal gain values and profiles are based on EN 16798-1. However, studies have shown that the real electricity use of plug load and lighting varies more smoothly than in the profiles of EN 16798-1 where zero occupancy outside working hours is assumed. This might result in sub-optimal building solutions due to inadequate building performance simulation input data. The aim of this work is to structure and analyse data from a total of 196 electricity meters in 4 large office buildings in Tallinn, Estonia. Typically, 3 to 8 electricity meters were installed per floor with the consumption coming mainly from plug loads and electric lighting. The data had been gathered between the years 2016-2020 with either 1 or 24 hour time steps, depending on the building and the electricity meter. 3 out of the 4 buildings had an average normalized energy usage slightly below the modelling value calculated according to EN16798-1. Some office spaces stood out with an abnormally high electricity consumption; however, the 24-hour distributions were fairly compact, meaning quite steady consumption patterns. When looking at the dispersion of energy consumption per 24h, averaged over all given offices in a building, no outliers stood out, either. This means that there are not many days when the average consumption and internal heat gains of all offices were simultaneously well below the mean. Additionally, major events like holidays and the COVID19-induced lockdown show up well on the graphs, but also planned changes in occupancy can be seen.

1 Introduction

Office buildings are well known to consume about 40% of the total energy share of the European building sector [1]. As the European Union (EU) has set long-term targets to reduce carbon emissions and energy consumption significantly, improving the energy-efficiency of office buildings is a priority. To this aim, researchers and designers are now focusing on structural improvement as well as on smart technologies, which can align building operation and occupants' needs.

Such alignment is a crucial characteristic of modern approaches: the heating system is now viewed as a means for temperature control rather than just emitting heat to rooms. It has been indeed demonstrated that internal and solar heat gains of intermittently operated buildings such as office buildings (OB) can cover the majority of heat losses [2].

This happens because heat gains from people, equipment and lighting as well as ventilation heat loss have a large impact on heat balance. The fluctuating heat gains and non-demand based ventilation operation make the thermal behaviour dynamic. This is not accounted for in the current design methods, whose conservative and simplistic approach of accounting heat gains results

in over-dimensioned and sub-optimally operated systems. On the contrary, the dynamics of heat balance and energy performance of a modern OB require a fairly complex analysis to be performed with advanced computational methods. Building performance simulations (BPS) provide a powerful tool in this sense already at the design stage: for instance, in Estonia it is mandatory to use dynamic (namely, hourly-based) BPS for calculating the Energy Performance Certificate (EPC) of commercial and residential buildings.

Unfortunately, even with simulations there often exists a sizeable difference between calculated and actual energy performance of buildings. Cali et al. [3] demonstrated that the consumed energy can be up to 3 times larger than the calculated estimates; occupants' behaviour was identified as one of the causes of the performance gap in addition to errors in installation and operation of the buildings. Several studies have therefore developed modelling strategies based on the monitored use of OBs, focusing either on occupancy [4], lighting [5] or plug loads/computers [6] measurements. It became immediately clear that the real electricity use of plug load and lighting varies more smoothly than in the profiles of default occupancy schedules building codes and standards, such as the EN 16798-1 [7], where zero occupancy outside working hours is assumed. More

* Corresponding author: andrea.ferrantelli@taltech.ee

often than not, plug loads and lighting consumption are indeed sizeable also outside occupied hours [8] [9]; significant variances between daily electricity uses of single occupants or office rooms do exist as well. It is thus necessary to track the OB's energy use during the entire 24 hours period, and to study how the measurements correlate with more sophisticated occupancy schedules. This knowledge can then be implemented into accurate BPS for guiding simulation-based design decisions. Reducing the size and cost of heating and cooling systems, simultaneously increasing their efficiency, will then lead towards a new generation of dynamic sizing methods for the heating and cooling of office buildings.

In this paper we lay down such experimental groundwork, which is critical for identifying the typical use of equipment and lighting in office buildings in order to develop methods for e.g. integration into building simulations in order to reduce energy use, improve tenants' comfort, building flexibility etc. The aim of this work is structuring and analysing data from a total of 196 electricity meters in 4 large office buildings in Tallinn, Estonia. Three to eight meters were installed per floor, monitoring plug loads and electric lighting. The data had been gathered between the years 2016-2020, with either 1 or 24 hour time steps depending on building and electricity meter.

Table 1 List of relevant properties for all the buildings.

Properties	A	B	C	D
No. of floors (total)	13	13	8	13
Monitored floors	7	12	8	12
Zones per floor	4	4	6	8
Total area monitored (m ²)	4938.3	8508.2	4052.0	13989.6
Customer service nr. of offices %	4 19%	2 11.8%	3 7.5%	5 4.2%
Administrative nr. of offices %	9 42.9%	11 64.7%	33 82.5%	107 89.9%
IT nr. of offices %	8 38.1%	4 23.5%	4 10%	7 5.9%
Total	21	17	40	119
Meas. points (total)	32	48	50	102
Outlier %	5.09	11.91	5.01	9.91
Resolution	0.001	0.001	1	1
Original data unit	kWh/h	kWh/h	kWh	kWh
Start date	2019-01-01	2019-01-01	2017-12-08	2018-01-05
End date	2019-12-31	2019-12-31	2020-03-07	2020-03-07

2 Methods

In this section we describe the datasets acquisition and structure, data preprocessing and methods of statistical analysis.

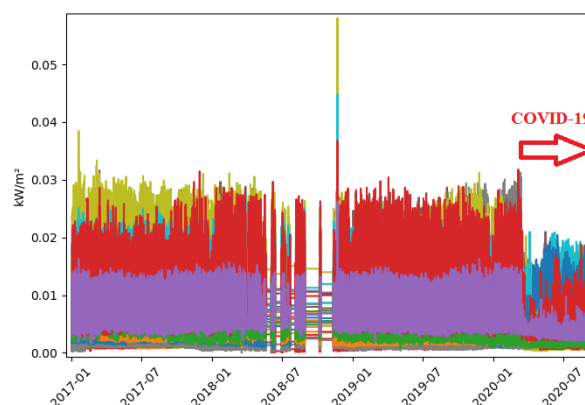


Figure 1 Raw data for Building A.

2.1 Datasets acquisition

This study is based on electricity consumption data acquired from four office buildings located in Tallinn, Estonia (**Table 1**).

Each floor of any building was divided into zones where electricity consumption was metered separately; most of the floors follow a standard layout, only the first and second floor have a larger area.

Each measurement point had three-phase electricity meters that were compatible with a 230/400 VAC voltage system. Measurements were performed with class B meters conforming to EN 50470-3, which had been installed during the construction of the building. The data acquired from the meters was stored in a building management system from which it could be downloaded into a CSV file.

The amount of data from each building ranged from 11 to 27 months. The time resolution of the data was preset by the building management system operator; two buildings displayed hourly data and the other two had daily data (**Table 1**).

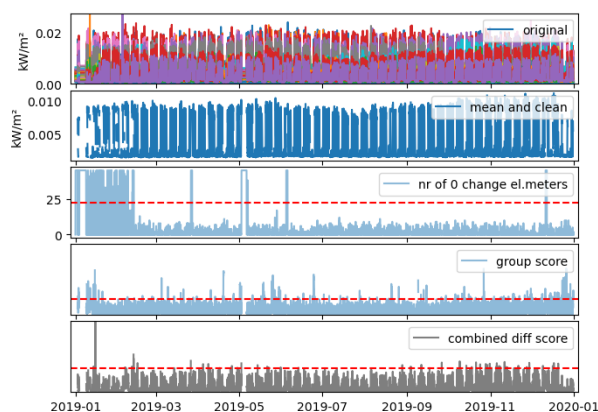


Figure 2 Raw data and outlier predictions for Building B.

2.2 Data preprocessing

2.2.1 Data format

A detailed overview of the building-specific parameters, including floor areas, measurement point counts and time ranges is given in **Table 1**. The original data came in two types of formats. For Buildings A and B, this was a non-cumulative series of hourly kWh consumption readings with 0.001 kWh resolution, for Buildings C and D it was a cumulative series of daily kWh readings with 1 kWh resolution. The data of Buildings C and D was then converted to a non-cumulative series (kWh/24h) by calculating the differences of two contiguous entries.

In order to give the finalized consumption values in units per square meter, the project documentation of the buildings was used to gain information about the serviced floor areas for each electricity meter. The official areas were given with 0.1m² precision, however, since there were some inconsistencies, we fixed the estimated error at 1m².

2.2.2 Data cleanup

As a first step, some periods of data were left out based on existing knowledge about building occupancy (see **Table 1**). A few electricity meters, which according to the project documentation were labelled as ordinary office meters, were also excluded, since their behaviour and power consumption were significantly different from a typical office meter's pattern, possibly monitoring the consumption of some mechanical equipment.

The COVID-19 impact could also be seen in the graphs as the power usage significantly dropped from March 2020 onwards (**Figure 1**), so the latest cut-off date for all buildings was set to March 7th, 2020.

After visualizing the time-series graphs of used power (kW), numerous other problematic time periods showed up, as in **Figure 1**. These were either affected by stuck readings or by abnormally high peaks. The reasons for such errors could have likely originated from the BMS (Building Management System). One possible explanation for such peaks is the accumulation of used energy while the BMS was shut down since most electricity meters do not log energy consumption with a timestamp. There is also a possibility of external interference in the measurements caused by electromagnetic compatibility issues or poor error mitigation inside the BMS. However, these eventualities have not been verified in the current scenario.

For further analysis of anomalous behaviour and outliers, an algorithm was developed to remove potentially bad data points. Some parameters were adjusted slightly for a couple of buildings for more optimal detection, but the general method is as follows, in the given order:

- If the reading of a single electricity meter at any given time is significantly different from the mean of all meters at that time, exclude the slice. This is necessary

to avoid losing outliers after the data has been averaged across a building.

- This step applies only to Buildings A and B (1h timestep). If more than half of the meters show static behaviour, exclude the slice. This is again necessary, since there appeared to be numerous small stops in the readings of individual electricity meters, additionally to the large, synchronized freezes mentioned before.

- Calculate the average consumption across a building at any given time.

- Group the averaged data by weekdays (and hours, if applicable) and exclude points where the value is further than 2 standard deviations of that group's mean.

- Create a combined score of first and second absolute differences of the series, where the second difference has a slightly higher weight. Exclude points where the combined score exceeds a threshold. This helps to remove smaller peaks and abrupt changes.

After visualizing the predictors with this method, some additional time periods stood out with poor behaviour, as can be seen on the left side of **Figure 2**, thus they were left out.

2.2.3 Conversions for distribution analysis (Buildings A and B)

The data for power consumption distribution analysis were given in units kWh/(24h·m²) for compatibility. This has already been achieved with Buildings C and D, but conversion was needed for Buildings A and B.

Since the data had been cleared of outliers, simply summing up the hourly readings of each day could have returned lower than actual results, due to missing values. However, excluding all days that have any missing data would result in a huge loss of data; to reduce the number of lost days, a linear forward interpolation of maximum four hours was thus applied before excluding days with any missing values.

2.2.4 Statistical data analysis

The cleaned-up data were processed with the software R [10] through various packages that allowed exploring distributions as well as performing normality and correlation tests. For Buildings A and B, the 1-hour data were used for daily and weekly analysis, while the monthly assessment used hourly data that were averages of all Mondays, Tuesdays etc. of that month. These correspond to the "average" or "representative" days that are addressed in the next Section.

3 Results

3.1 Boxplots of daily consumption

3.1.1 General

The plots in **Figure 3** to **Figure 6** show distributions of datapoints in kWh/(24h·m²) of each measuring point

in a building. The whiskers of the boxplot are drawn at 5th and 95th percentiles. Green triangles represent arithmetic means and green lines represent medians.

The red, dashed horizontal line represents the reference value used for modelling energy consumption of office buildings, calculated according to EN16798-1 [7]. The value is 0.1089 kWh/(24h·m²), which assumes power consumption of 0.018 kW/m² (0.006 for lights and 0.012 for equipment) at an average usage level of 55%, over an 11h period in a day. The blue horizontal line represents the calculated average consumption of the selected offices in a building.

3.1.2 Analysis

For Buildings B, C and D the average falls only slightly below the reference value, but for Building A the average consumption is significantly higher. It seems that most of the monitored offices in Building A have an average consumption above the reference value, so the high average is not caused by any outstanding offices, rather from a general behaviour of the occupants. **Table 1** displays more administrative than IT offices, however the high consumption should not be related to pc use only. We have no info about employee number either, so no correlations can be generated between user profiles and consumption patterns.

However, the two zones with the highest consumption in Building C are known to be dentist offices. On the far right, a column called “AVG” shows the distribution of average daily consumption values of all the measurement points combined. The variance is quite small, compared to the variance of all other offices in that given building.

This means that there are no large, synchronized swings in the building’s total consumption, which can also be seen in **Figure 2** in the graph “mean and clean”.

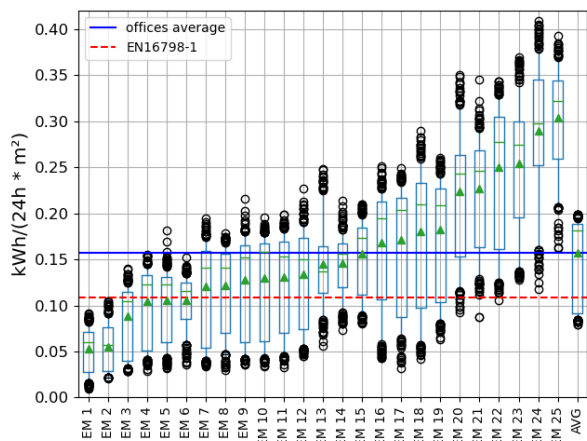


Figure 3 Boxplots for Building A.

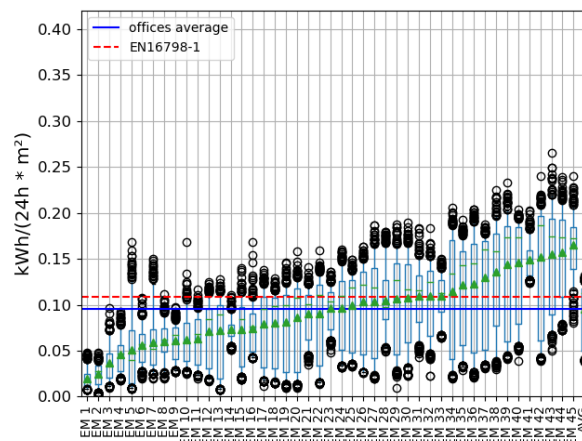


Figure 4 Boxplots for Building B.

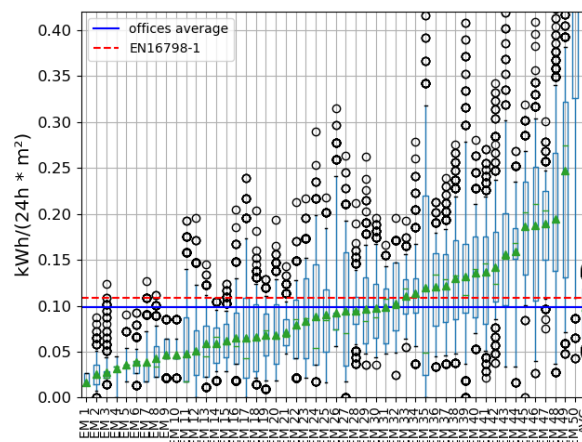


Figure 5 Boxplots for Building C.

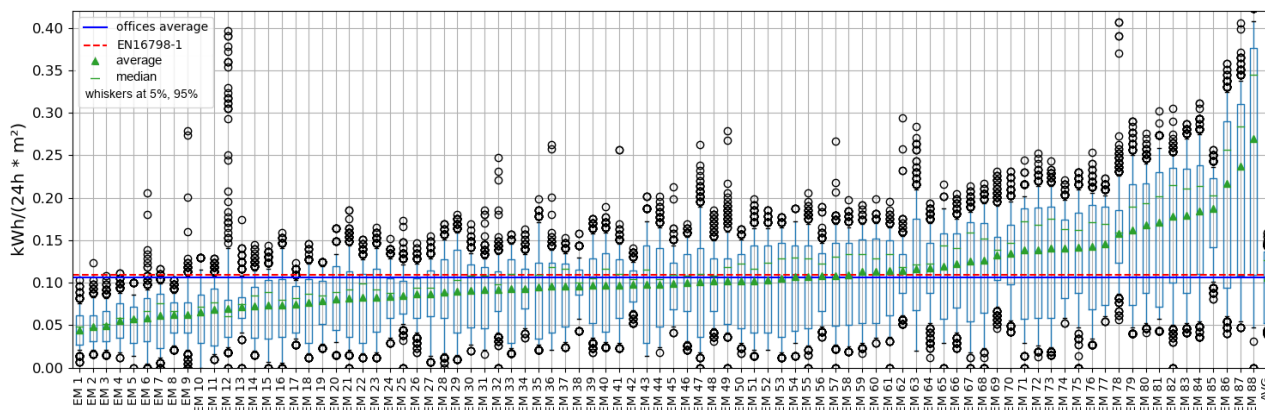


Figure 6 Boxplots for Building D.

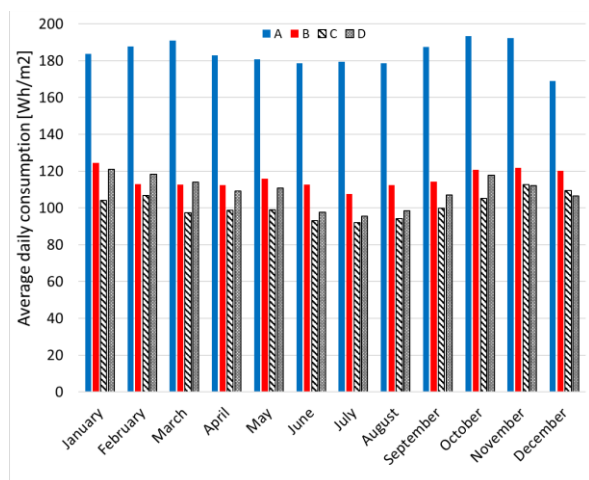


Figure 7 Cumulative daily power consumption, 2019 monthly breakdown for all buildings.

3.2 Monthly analysis and seasonal variations

For each of the four buildings, a monthly breakdown of weekday cumulative consumption [Wh/m²] for the year 2019 was computed. This was obtained, for buildings A and B, by adding all the average hourly values; for buildings C and D cumulative daily values were averaged (24h time step for the data). The result is plotted in **Figure 7** for each building.

Considering the full interval 2016-2020, small differences among the years do exist, whilst the overall pattern does not change qualitatively. Consumption is higher in the Autumn and Spring months, not during the winter as generally expected. For each case, we found very little correlation between climate and tenants consumption: let us remind that only plug load and lighting consumption were monitored, not heating.

Let us consider Building A as an example: a Pearson correlation test returned 0.999 for January versus June, showing a high correlation between winter and summer months. This was confirmed by a Kendall test (more sensitive than the Pearson test) as well.

Furthermore, the same test provided -0.69 and -0.68 for tenants consumption versus, respectively, measured sunshine duration and daily external temperature, which is indicative of a weak correlation. For 2019, the largest

consumption was recorded in March, followed by October and November.

It is interesting to investigate the role of sunshine duration more into detail, since our data addressed both equipment plug load and lighting. Since sunshine duration accounts for cloud coverage, differently from daylight hours, it can influence switching lights on and off. A plot of daily average power consumption in function of measured monthly sunshine hours is given in **Figure 8** for Buildings A and B, and in **Figure 9** for C and D. For Building B we used January 2020 data, as the January 2019 data were not sufficient for the statistics. Remarkably, January 2020 was as sunny as March 2019, namely over 3 times sunnier than January 2019. It was also much warmer, with average T=3C versus -3C in 2019. Yet, its average daily consumption was 10% larger than February 2019 (T=1C) and March 2019 (T=2C), confirming the importance of occupancy schedules.

At the building level, **Table 2** features the correlation matrix of 2019 monthly consumption for the four datasets. Building A is fairly uncorrelated from the others, consistently with e.g. **Figure 7**, while B and C seem to be slightly more comparable. Although the fact that A and B have 1hr and C and D have 24hr data hinders any speculation about occupancy patterns, the low overall correlation mirrors the absence of a common climate-induced seasonality in the data.

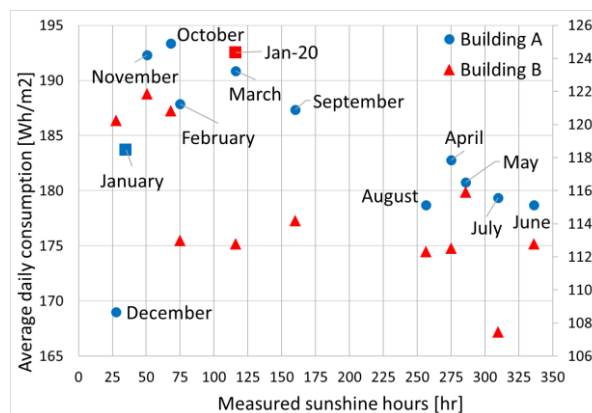


Figure 8 Building A (dots, left axis) and B (triangles, right axis) - Cumulative daily power vs monthly sunshine hours.

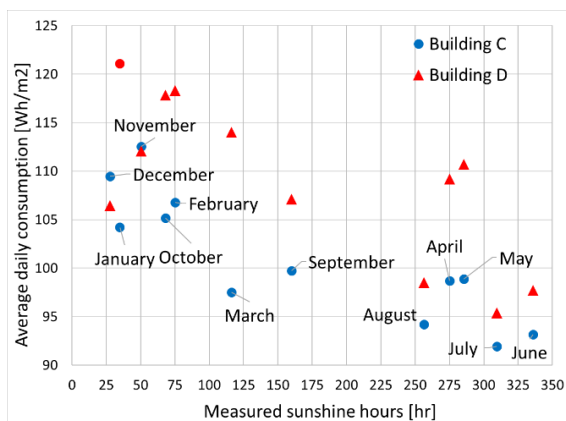


Figure 9 Building C (dots) and D (triangles) - Cumulative daily power consumption versus monthly sunshine hours.

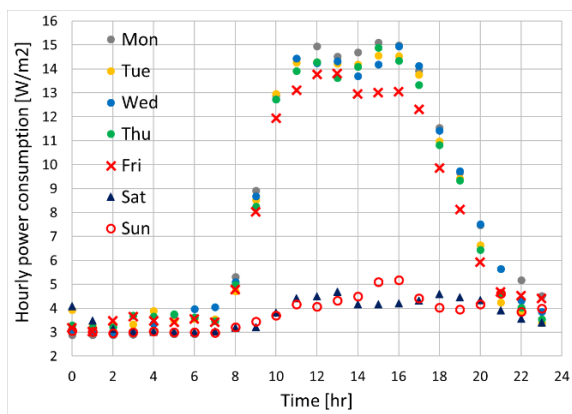


Figure 10 Building A - Tenants' consumption for January 2017, representative week.

Table 2 Correlation matrix for the four datasets, 2019.

	A	B	C	D
A	1.000	0.273	0.333	0.455
B	0.273	1.000	0.636	0.515
C	0.333	0.636	1.000	0.455
D	0.455	0.515	0.455	1.000

3.3 Daily analysis

3.3.1 General considerations

The tenants' electric consumption is illustrated for a representative January 2017 week in **Figure 10**. Notice the sharp decrease at lunch break and the lower consumption for Fridays, as expected.

Qualitatively, the weekday curves do not differ much between winter and summer months, confirming the high correlation already discussed. This agrees with the data distribution, which is sharply bimodal with the two peaks at the histogram extremes for each month.

We recall that Buildings C and D provided only with 24h data, therefore it was not possible to investigate the hourly breakdown as in **Figure 10**. This section will therefore discuss our findings only for A and B. Considering a specific day with expected full occupancy, we chose a central Wednesday in January

2017. The daily consumption reflects our results for monthly averages: normality is confirmed by QQ (Quantile-Quantile) plots and a Cullen and Frey plot, while histograms show a clear bimodal pattern with modes at the extremes. The statistical parameters of the distribution are an estimated standard deviation (SD) of 4.797, a skewness of 0.288 and a 1.219 kurtosis. The large SD and low kurtosis signify that, despite the substantial data cleanup described in Section 2.2.2, we are still in the presence of outliers, as illustrated in Section 3.1.

3.3.2 Energy consumption prediction formulas for building performance simulations (BPS)

Aiming at using our measurements for implementation into BPS, we generated prediction formulas of energy consumption by adapting to our dataset a bottom-up method that was introduced for domestic hot water data in [11] and then applied to buildings' energy consumption in [12]. The case of [12] addressed a much larger building, with relatively small variances in the hourly consumption profiles for different weeks and months, so it was possible to identify a *unique* representative day whose consumption could be correlated to other days, to cover a full year.

On the other hand, for Building A (and even more for B) too many days had very different profiles, requiring a less simplistic approach. For instance, July 2019 showed the cumulative consumption of the most correlated Monday to be equal to that of the average Monday, while for Tuesdays the difference was remarkable, 5.62%. Preferring an average day to a specific day as representative was therefore more suitable.

Since hour-by-hour prediction in this case is not reliable, we focused on predicting the cumulative energy consumption with the lowest error possible; we also wished to keep smooth interpolation curves to avoid too biased predictions. The procedure followed these steps:

1. The cumulative consumption of *average* days for each month is split into four groups: Mon to Thu (WD), Friday, Saturday and Sunday. The value that is closest to the average is called E_{wd} , E_{Fri} , E_{Sat} , E_{Sun} . For Building B, the corresponding days were February Wednesday, March Friday, June Saturday and September Sunday.

2. Interpolate each of these four reference days and obtain the fit formulas $E_{fit,WD}(t)$ etc. These are our "structural formulas" according to the terminology of [12]. Consumption for a random day can now be predicted by using linear correlations with the formulas

$$E(t) = A * E_{fit,WD}(t) + B, \quad [W/m2] \quad (1)$$

for weekdays and

$$E_{i,m}(t) = E_{fit,i}(t) * R_{i,m} \quad [W/m2] \quad (2)$$

for Fridays and the weekend, where $i=Fri, Sat, Sun$ and $m=1, \dots, 12$. The coefficients A and B are computed by correlating each month with the one corresponding to

the structural formula. For Building B, February holds for weekdays, with coefficients listed in **Table 3**.

Table 3 Monthly correlation formulas for Building B.

Month	A, B (Mon-Thu)	R ²
January	1.0891, 0.0545	0.9739
February	1,0	1
March	0.957, 0.1653	0.9965
April	0.9406, 0.2385	0.991
May	0.9925, 0.1338	0.9963
June	0.945, 0.2415	0.9923
July	0.8788, 0.3439	0.9929
August	0.9478, 0.2147	0.9932
September	0.9785, 0.1522	0.9938
October	1.0452, 0.1172	0.9772
November	1.0753, 0.0216	0.9719
December	1.0374, 0.1425	0.9803

For Friday and the weekend, one should instead use Eq.(2), where

$$R_{i,m} = E_{i,m}/E_{fit,i} \quad (3)$$

(i=Fri, Sat, Sun) is a non-dimensional coefficient that is computed for each case, given the cumulative energy of the average desired day $E_{i,m}$ (say, for January Friday) and that of the fitted day $E_{fit,i}$. The $R_{i,m}$ coefficients for Building B are listed in **Table 4**.

Table 4 $R_{i,m}$ coefficients for Building B.

Month	Fri	Sat	Sun
1	1.074	1.056	1.079
2	0.973	0.935	0.925
3	1.000	0.984	0.945
4	0.992	0.976	0.947
5	1.019	0.960	1.005
6	0.984	1.000	0.984
7	0.928	0.960	0.959
8	0.974	0.978	0.964
9	0.985	0.978	1.000
10	1.041	1.031	1.021
11	1.044	1.031	1.029
12	1.024	1.042	1.056

Using the above method, we could predict the consumption with small deviations from measurements. The errors ranged from 0.96% to 3% for Mon-Thu, 0.94% to 4.31% for Friday, 0.15% to 3% for weekends.

4 Discussion

During analysis, numerous concerns arose regarding the actual reliability of the data. Inconsistencies in the project documentation, occasionally over-dimensioned

electricity meters, numerous logging problems with the BMS and major occupancy changes resulted in a dataset that was far from ideal.

However, it can be shown that computing the error propagation for weighted averages of all electricity meters per unit floor area resulted in negligible final error bounds (0.004 W/m² on the average). This suggests that installing more meters could produce more accurate data, giving an advantage over measuring everything at the building level only. Additionally, having more meters would allow excluding undesirable zones, which will very likely be present, as well as leaving the ability to distinguish between different types of consumption. It is indeed well known that diverse space-use typologies (distinguished by a combination of tenants' tasks and time-based occupancy) generate a variety of daily consumption profiles, see e.g. [9]; a whole-building zonal analysis would thus allow tailoring the HVAC design to these diverse needs.

One of the data features that could be learned from the box plots comes from the average daily consumption values of all the measurement points combined, which has a small relative variance for that given building. This means indeed that there are no large, synchronized swings in the building's total consumption, which can also be seen in **Figure 2** in the graph "mean and clean".

However, further analysis can be done about the upper, 95th percentile values of individual offices, especially for buildings where hourly data is also available. This could give more information about local peak loads for dimensioning mechanical equipment, as well as finding different correlations.

Our efforts in measuring electricity locally instead of per building showed that it is advisable to invest in measuring electricity locally, rather than being content with measurements at building level, for two reasons. First, the variance induced by diverse types of offices is substantial; this important information disappears if data are aggregated for the whole building. Secondly, our statistical analysis of monthly and daily patterns showed a lower impact of climate and irradiation hours than expected, illustrating the predominant role of occupancy that strongly depends on the specific office typology.

The analysis of monthly consumption brought forward some interesting non-trivial features. First of all, the high correlation between daily profiles during winter and summer months is a signal of a recursive pattern that is not influenced by added sunshine hours.

Figure 8 and **Figure 9** also illustrate that although some correlation with sunshine duration does exist, this is seemingly dominated by the plug load. In **Figure 8**, March dominated over January and February, and April and May over August (Building A). This is common to all the buildings here studied (see also **Table 2**), suggesting a central role of occupancy consistently with [12] and underlining a necessity to address its impact thoroughly.

The daily analysis showed that, although the absence of a standard energy profile was problematic for an hour-by-hour prediction, by focusing instead on the cumulative consumption we managed to establish a procedure that allows implementation into BPS with

good accuracy. The basic idea was that the inconsistencies tend to compensate each other, giving room to some unavoidable tolerance.

5 Conclusions

In this paper we have investigated plug loads and lighting consumption data of four office buildings in Tallinn, Estonia, over a four years period. Data acquisition and preprocessing of some very problematic measurements were discussed into detail, together with a simple, yet effective prediction method for application into BPS towards energy estimates.

Among the other results, we have demonstrated that it is preferable to install more meters rather than measuring everything at the building level, for increased accuracy and for keeping relevant information. Consistently with previous studies, we also found that occupancy patterns are central in determining the electricity consumption.

The definition of a typical office building should therefore be discussed (e.g., IT and administrative work can be quite different in terms of energy use intensity), occupants' density (measured or estimated?), installed plug loads and lighting power etc. These could all provide useful information in order to shift the focus of energy performance research, in order to consider the actual energy use.

We wish to remark that the amount of data analysed is quite remarkable by itself: whilst not providing a fully exhaustive overview of energy consumption in non-residential buildings, it is larger than the average datasets that appear in this type of studies.

Overall, the gathered information has a number of applications on different levels, from tailored predictions aimed at renovations, to refinement of applied predictive modelling strategies, to classification and benchmarking of building energy consumption.

Considering our findings and the above improvements, this study and its developments have the potential to contribute to future calculations of energy performance estimation in office buildings. And even if after COVID-19 we may never go back to the old way of office use, our dataset finds formal application in predictive modelling strategies. It also constitutes a good basis for energy consumption benchmarking, as it provides a baseline upon which future optimisation strategies based on new working styles can be compared.

The authors acknowledge support from the European Regional Development Fund via the Estonian Centre of Excellence in Zero Energy and Resource Efficient Smart Buildings and Districts ZEBE, grant 2014-2020.4.01.15-0016, from the Estonian Research Council through the grant PSG409, from the programme Mobilitas Pluss (Grant No. 2014-2020.4.01.16-0024, MOBTP88) and from the European Commission through the H2020 project FinEst Twins (grant No. 856602). They are also grateful to Andrei Engels for compiling the raw data files.

References

- [1] European Commission, "Energy Performance of Buildings," 2020. [Online]. Available: https://ec.europa.eu/energy/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en. [Accessed 9 December 2020].
- [2] M. Thalfeldt, J. Kurnitski and A. Mikola, "Nearly zero energy office building without conventional heating," *Est. J. Eng.*, vol. 19, no. 4, 2013.
- [3] D. Cali, T. Osterhage, R. Streblov and D. Müller, "Energy performance gap in refurbished German dwellings: Lesson learned from a field test," *Energy and Buildings*, vol. 127, pp. 1146-1158, 2016.
- [4] J. Zhao, B. Lasternas, K. P. Lam, R. Yun and V. Loftness, "Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining," *Energy and Buildings*, vol. 82, pp. 341-355, 2014.
- [5] S. Gilani and W. O'Brien, "A preliminary study of occupants' use of manual lighting controls in private offices: A case study," *Energy and Buildings*, vol. 159, pp. 572-586, 2018.
- [6] R. M. Tetlow, C. van Dronkelaar, C. P. Beaman, A. A. Elmualim and K. Couling, "Identifying behavioural predictors of small power electricity consumption in office buildings," *Building and Environment*, vol. 92, pp. 75-85, 2015.
- [7] E. 16798-1, "Energy performance of buildings - Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings," CEN, Brussels, Belgium, 2019.
- [8] A. C. Menezes, A. Cripps, R. A. Buswell, J. Wright and D. Bouchlaghem, "Estimating the energy consumption and power demand of small power equipment in office buildings," *Energy and Buildings*, vol. 75, pp. 199-209, 2014.
- [9] A. Mahdavi, F. Tahmasebi and M. Kayalar, "Prediction of plug loads in office buildings: Simplified and probabilistic methods," *Energy and Buildings*, vol. 129, pp. 322-329, 2016.
- [10] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, 2020.
- [11] A. Ferrantelli, K. Ahmed, P. Pylsy and J. Kurnitski, "Analytical modelling and prediction formulas for domestic hot water consumption in residential Finnish apartments," *Energy and Buildings*, vol. 143, pp. 53-60, 2017.
- [12] A. Ferrantelli, H. Kuivjõgi, J. Kurnitski and M. Thalfeldt, "Office Building Tenants' Electricity Use Model for Building Performance Simulations," *Energies*, vol. 13, p. 5541, 2020.